California Energy Commission Award No. 500-10-052

National Lab Buildings Energy Efficiency Research Projects

LBNL EF87EE

Task 2.11: Improved Audio-Video Efficiency Through Inter-Device Control

Deliverable: Summary of Analysis of Communication Link Technologies

Bruce Nordman

September 21, 2012

Background

This document is the second deliverable for *Task 2.11, Improved Audio-Video Efficiency Through Inter-Device Control*. The research conducted in this task comprises one element of the National Lab Buildings Energy Efficiency Research Projects, CEC Award No. 500-10-052. The objective of this task is to save substantial amounts of energy by creating a technology standard for how inter-connected audio/video (A/V) devices manage their own power state. The goal is that the technology standard will be incorporated into future products and communication standards.

This report assesses common communications technologies and data transport found in A/V systems, both legacy analog and digital technologies. Implications for power management and energy savings are considered and summarized.

This document has eight sections, including this. The Introduction presents and defines concepts fundamental to the analysis. The third, fourth, and fifth sections review Analog, Digital non-IP, and Digital IP Technologies. AV Streams are described in the sixth section. The seventh section - Control Paradigms - considers the evolution of A/V controls. The document ends with a Conclusions section.

Introduction

The purpose of this report is to assess common communications technologies that transport data for audio/video systems — those usually centered around televisions. The information of interest is principally that which relates to how the power or functional state of devices is communicated over each link:

A **link** is a communications connection from one device to a second device. A data link is between two devices with an electrical connection. A network link is a logical connection that may require one or many data links.

The **power state** is whether the device is on, asleep, or off. Power state can be changed manually (by a mechanical switch or a device's remote control), through automatic internal operation of the device, or by command from a connected device. We consider that a device that is off can only be brought out of that state by an explicit power command, whereas a device that is asleep can wake up by some ordinary interaction: by the device itself, via a remote control, or due to information exchanged on a network connection.

The **functional state** includes the source to which a device is currently listening to (e.g. the input channel used by a TV or A/V receiver). It can also include the type of information involved (e.g. only audio, only video, or (usually) both). Functional state could also include whether the data is being further transmitted to another device (e.g. a TV mirroring its display over a network to another TV elsewhere), and locally acquired data such as occupancy.

Power state and functional state are entangled with each other, in internal device operation, in inter-device communication, and in user experience. Power state changes affect functional state, and functional state changes sometimes affect power state. The key is to do so in a way that is clear and works well for user experience (convenience) and energy use.

Analog Technologies

Audio/Video devices used only on analog connections for many decades, with significant use of digital methods only in the last ten years. We can expect the use of analog interfaces to continue for a some time, particularly in speakers, including earphones. Some analog connections are used for only audio, some only video, and some both. In general, analog connections do not allow for communications other than the actual content being transmitted. Also, no communications are possible from the receiver to the sender, with three exceptions: the absence of any signal at all; a "blank" signal (no sound for audio, and an entirely black image for video); and cable insertion (mechanical presence in the jack).

Audio

The two common analog audio-only technologies are line level audio (usually stereo), and powered audio (for driving speakers). A variety of connector types can be used for line level audio, though "RCA plugs" are most common in residential systems. On a TV, these are the red and white plugs that accompany composite or component video connections.

There are products on the market that are able to detect no audio signal present, and then power down. The most common example is the powered subwoofer; they are usually remote from the main equipment rack, are often inconveniently located, and usually the only user interface is a power switch (and may not have a power switch at all). In these cases, it is usually necessary or desirable for the device to be asleep, and not off, when powered down, so as to be able to sense when a signal reappears and then wake up.

Video

Video transmission requires sophisticated timing to divide up time into frames, into scan lines, and then into the data along each scan line, and also to encode color information. The three basic mechanisms for video transmission in the U.S. are Component, Composite, and S-Video. For our purposes, these have the same capabilities, namely, that only the absence of a signal or a totally black signal can be detected by the receiving device. Some devices can detect the absence of a signal — a common example of this is projectors, which cycle through their inputs looking for an active feed. Most computer monitors will go to sleep in the absence of a signal (this is distinct from the sleep signaling in the communications standard itself), and many TVs will power down to off if no signal is present for an extended period of time.

In Europe, the SCART (a French acronym) connector was widely used for televisions until the introduction of HDMI. SCART also includes audio (both directions), and various types of video. A digital communications channel was added to SCART in the mid 1990s that included functional commands such as changing channels and volume, and also includes power commands. SCART was adapted to become the CEC part of the HDMI standard (see below).

The VESA (vesa.org) standard for computer displays (commonly called VGA for the first resolution supported) encodes analog video information. It does not include audio, but it does include a digital data line for negotiation resolutions and other features. The VESA standard for Display Power Management Signaling (VESA, 1993) specifies how a display should go to sleep based on the status of the horizontal and vertical synchronization signals. Some TVs and A/V receivers include VESA inputs which allow them to display a computer's output. VESA also hosts the specification for DisplayPort (see below) but we use the VESA term to refer only to the analog interface.

While it would be possible to evolve the signaling in SCART and VESA interfaces using their digital lines, improvements are unlikely to occur. Use of SCART and VESA are declining, so upgrades would be of marginal use going forward, since both connected devices would also need improvements to support the advanced capability.

Control

There is one analog control technology (as opposed to a content mechanism) in A/V products, namely the 12V trigger. This is a cable used so that one device can turn other devices on or off as a toggle. Since it is a toggle, not a deterministic setting, this works best if the controlling device is the only one that controls the power state of the controlled device. This technology is usually used only in professionally-installed systems.

Summary

Analog technologies raise a set of fundamental questions for power control. Specifically, whether a device can know if: there is a cable attached, there is a device at the other end of the cable, that device is on, or it is "listening" to that channel. Even if some of these questions could be answered electrically, it is not known whether the process would be reliable, and whether manufacturers could be compelled to develop and commonly include this feature. Some devices sold today detect mechanical cable insertion — for example, computers that switch off internal speakers when a cable is inserted into the earphone (line level audio) jack — but this is mechanical detection, not electrical sensing. The cable detection mechanism can determine that there is no device to send to, but presence of a cable does not assure that there is a receiving device nor that any connected device is powered on.

In summary, in analog connections a receiving device can usefully detect signal presence and respond by changing power and/or functional state, but a sending device cannot. Cable detection may have a limited use. Analog connections should not drive the design of our overall architecture for inter-device power control, but should be taken into account once it has been established.

Digital non-IP Technologies

There are a large and rising number of digital technologies involved in distributing audio and video. In this section we cover all technologies except those that are based on the Internet Protocol (IP) — they are covered in the following section. While non-IP technologies have "layers" of communication internally, they typically specify all of them in a single standard, unlike the sophisticated layering used in IP technologies.

HDMI

The High Definition Multimedia Interface (HDMI, hdmi.org) enables more than a data link but less than a network link. HDMI (HDMI, 2004) devices create a "tree" of connectivity, with the root at a TV, leaves at source devices, and intermediate devices between that can select among several devices. The simplest example of this system in composed of just two devices — a source and a sink. The most common intermediate device is an A/V receiver.

Capabilities

Many details of HDMI are interesting and useful but do not affect power control so are not covered here. These include management of different video resolutions, content protection, and audio options. The mechanism for transporting the actual video data is essentially the same as the DVI standard (see below), which is uncompressed and unencrypted. HDMI includes a copy protection mechanism which has been critical for its industry acceptance.

HDMI is plagued by multiple versions with incremental additional functionality, and some incomplete, inconsistent, or otherwise problematic implementations. This can cause confusion and frustration for those installing systems, but in many cases, it works just fine. Problems are most likely to occur when the technology or products are "stressed" as with large collections of devices, long cable runs, or use of uncommon features of the standard.

HDMI is a collection of data transfer mechanisms. The primary mechanism is the routing of video data, and audio is also encoded on this same path. A second path is for routing 100 Mb/s Ethernet frames (for general use, not necessarily related to the video content), and a third path sends audio in the reverse direction (for example when the TV is the source of content and is passing it back, upstream, to the receiver for sending to the speakers)[1]. HDMI also provides 0.25 W of power, which is useful to power active adapters that may be present, and enables a "hot plug detect" mechanism to signal when there is a device at the other end of the line. Finally, there is a communications line for describing display capabilities (so correct resolutions are transmitted), and the Consumer Electronics Control (CEC) line described below.

Addressing / Topology

To enable communication among the devices that HDMI interconnects, HDMI assigns "addresses" which reflect the device's place in the connectivity tree and their function. As long as the device location does not change, the address should not change. The CEC channel is the mechanism used to map the topology of the "network" and to label each device. Each device has a "physical address," which describes the connectivity used to get to it from the TV (which is the root of the tree), and a unique logical address which describes the kind of device (when there is more than one, they are numbered). Devices are expected to participate in the CEC protocol whether they are fully on or not. This has a practical purpose — if you want to send a command to a device to power up, it has to be on the network, be known to the sending device, and receive the power up command. The CEC bus is a single communications pool shared by all connected devices, so that one device can broadcast messages to all devices, or send a message to any other single device.

In an HDMI system, there is a single stream of A/V data that flows from a source device, through zero to two intermediate devices, and ultimately to a TV that is a sink device. While this is the archetypal system, if a recording device is in the middle of the stream, it can become a sink if the downstream devices are not using its stream (the TV in the setup may be displaying content from a different source or powered down). In addition, there are more complex systems (usually only installed by professionals) with "HDMI switches" that can route multiple HDMI inputs to multiple HDMI outputs. These situations break the foundational model of HDMI and do or will require a more sophisticated approach.

Control

The Consumer Electronics Control (CEC) line implements a command language derived from that used in the SCART interface in Europe. CEC is necessarily involved in power control as it includes power commands and functional commands that implicitly do or may affect power state. Unfortunately, many parts of CEC are optional to implement, and provision is made for extensions specific to a given manufacturer, so that interoperability among systems for CEC capability is limited in most cases. In addition, it is optional to implement the higher-level CEC commands at all. The technology and requirements can be expected to evolve in the near term.

The HDMI organization codified its understanding of CEC in (HDMI, 2009). This references HDMI version 1.3 (2006). Since version 1.4b (2011), the HDMI specification became confidential to companies participating in the HDMI organization, so this discussion does not reflect any changes that may have been made since.

CEC is a combination of address management, internal, and external messages. Devices use internal messages to discover or announce what is going on, and external messages, that implement user requests. The CEC documents refer mostly to "standby," but sometimes use "off", apparently as a synonym. Even if a device is disconnected from power, it might be able to respond to CEC commands by using power from other connected devices, or provided over the cable. However when disconnected from the cable, the device would not be able to turn

[1] An amplifier may be active in generating audio even if the video for a TV is not passing through it — the TV may generate the video internally (a tuner or from an Internet connection) or from a non-HDMI input. In such a case, the audio is sent to the amplifier via the Audio Return Channel feature of HDMI.

fully on as there is no source of power to use. So, for practical purposes, the possible states are On, Standby, and Disconnected. These could be readily mapped to On, Sleep, and Off (with many devices lacking a state where they are plugged-in and off). Some receivers support pass-through of HDMI signals when off (in "standby"). It is notable that while HDMI does not use the term "sleep," HDMI 1.3 refers to a device "waking up."

The CEC protocol is cognizant of devices being on or not, and has commands for powering up and down, and mandates or advice as to when a device should do so on its own. CEC also recognizes that some commands should be ignored, such as a command to power down sent to a device which is actively recording something.

CEC enables passing of remote control commands from one device to another (e.g. so that a volume change can be passed from a TV remote control to the A/V Receiver that is powering the speakers). Some CEC commands are particularly relevant to power control. One is "One Touch Play", which causes: a source device, any intermediates and the TV to power up if needed; any intermediates and the TV to select the right input; and the source device to initiate sending content. A second CEC command is "System Standby" which in principle commands all connected devices to power down, though it is noted that any device involved in an ongoing recording should ignore the command.

With HDMI, a source device can assert to the rest of the system that it should now be the source for the TV, or the TV can do the same. That is, any device can be authoritative.

HDMI commands are a mixture of commands and confirmations. For example, a TV may command a device like a DVD player to become the video source, and then the DVD player will announce that it has become the source.

Tunneling HDMI

HDMI signals can be sent over a variety of intermediate physical layers, typically by inserting two devices that convert the signal from HDMI to the other format, and then back again at the other end. When done this way, it need not be standardized and (in theory) doesn't affect the signal. One form that is actually built into some devices is HDBaseT (see below).

WHDI

The Wireless Home Digital Interface (WHDI) technology (whdi.org) essentially puts a wireless link in the middle of an HDMI cable. Neither the sending nor receiving HDMI device needs to be aware that the wireless link is present — it is intended to be purely transparent. The initial deployment context is to have a WHDI adapter box[2] adjacent to each device.

---

[2] An example WHDI product is from IOGEAR (Wireless HD 3D Digital Kit). This also provides an HDMI output from the transmitter for a local TV. The signal is then "mirrored" to a remote TV. The devices get the characteristics for each TV to select the best content that both can accept, particularly for the resolution, as the sender ensures that it sends a signal that the receiver can understand and use, and it thinks it is sending to only a single display. In addition, the transmitter adapter can take input from either of two HDMI ports (selected with a remote control).

The IOGEAR product has power buttons on each of the adapter boxes as well as on the remote control. While it is certainly possible to manually control the power, presumably most people will simply leave both boxes on all the time. The power consumption of these is likely low, but this is the type of device that should have an explicit sleep mode to use less power.

The IOGEAR device takes "up to 20 seconds" to establish a connection. This delay could lead to many people leaving the devices on all the time to avoid the delay. On the other hand, they do have a "power saving mode" which appears to be simply an auto-off function.

It is possible to build WHDI hardware directly into a sending or receiving device to avoid needing the external adapter (several major TV manufacturers are formal "Adopters" of the technology). One can buy USB dongles to send a computer's display output via WHDI.

DisplayPort

DisplayPort (vesa.org) is the current high-performance link between a computer and a monitor (VESA, 2008). It has extensive capabilities to move data for many and high resolutions, multiple displays, etc. It has not gained traction for ordinary audio/video devices, and as HDMI is already common on A/V devices, there is little incentive for manufacturers to also include DisplayPort.

DisplayPort is more flexible and extensible than HDMI, so for niche applications (e.g. signage displays) that have non-standard requirements (e.g. more displays, longer distances, lower frame rates, etc.) it can be attractive. For example, up to 63 separate video streams can be transmitted.

DisplayPort has a separate control channel (AUX CH), a hot-plug detect mechanism similar to HDMI, and a display wake event which can be passed up to a source device via the AUX channel. This channel includes a mechanism by which a source can set a sink interface to Sleep mode, much as Display Power Management Signaling does for the traditional VESA interface. The device may then go to sleep, or may not if it is still performing other functions. The interface must wake and reply within 1 millisecond if requested. The AUX CH channel is a simple data link, unlike the bus model of HDMI. DisplayPort has the concept of a Stream Policy Maker that supervises stream creation and termination.

The Thunderbolt (Intel, 2012) interface is a superset of DisplayPort, also routing high-volume generic data (PCI-Express) in addition to the display information, and also providing up to 10 W of power (Apple in particular has embraced Thunderbolt). Thunderbolt does not appear to depart from DisplayPort for our purposes, except that someone might choose to send IP packets over the link instead (and then it would raise the issues in the IP section below).

DVI

The Digital Visual Interface (DVI) transmits video only, and the technology for doing the bulk data transmission is actually the basis of HDMI, though HDMI adds a variety of other functions and wires. DVI is falling out of use, and had its main application in connections between computers and monitors.

WiHD

Wireless HD (wirelesshd.org) is a local networking technology in the same high frequency band as WiGig (IEEE 802.11ad — see below under Wi-Fi). It accomplishes much the same goals, but uses non-Wi-Fi mechanisms. It does have a control mechanism similar to HDMI/CEC, called Device Control (pat of Audio video control — AV/C) including mechanisms to change power state. It also has a mechanism to tunnel HDMI signals such as CEC. These seem to be mostly duplicative.

The system also forwards infrared remote control signals from the receiver (in the remote location) back to the transmitter, and then on to devices that its local "IR blaster" is connected to. This enables remote commands (e.g. channel change) to be registered locally, but not local commands (e.g. mute the TV) to be executed remotely. This shows both the challenges introduced by multiple inputs/outputs and the benefits of true network connectivity (where sharing information in both directions would be easy).

WiSA

WiSA is the Wireless Speaker & Audio Association (wisaassociation.org). Not surprisingly, the WiSA technology only transmits audio. The transmitter is either a separate device that takes in the audio from the A/V receiver (or other device), or is built into an A/V device with some other purpose. The standard enables auto-configuration of the speakers, so that they can locate the spatial differences between them to create a two-dimensional map. The listener then enjoys speaker delays adjusted to optimize the audio experience. When a transmitter stops sending data, a speaker can then go to sleep. The standard does not require a power-saving sleep mode, but does not seem to pose any barriers to doing so, and any wake-up latency must be within one millisecond. The standard includes two-way communication with each speaker.

IEEE 1394/FireWire/i.Link

This standard (completed in 1995) is formally a link technology like USB but can be used to route data over multiple links. It has often been used for audio/video data, particularly with video cameras, but has been losing market share rapidly and is unlikely to be relevant for future products.

S/PDIF

This is an interface for transmitting audio data only over either coaxial cable or optical fiber. It was developed in the 1980s when the introduction of compact discs meant that digitally encoded audio was (in principle) commonly available. No control signals are utilized, and the data transmission is only one-way, so that it functions essentially like an analog interface would: the data are either present, or not, and the sender does not know if there is a device receiving the data.

MHL

The Mobile High-definition Link (MHL, mhlconsortium.org) is a technology that uses existing communications interfaces (e.g. USB) and adds functionality to them to enable streaming of high definition video. The motivation is to support mobile devices that have very limited connector options because they are small. This is a wired point-to-point connection for *ad hoc* connectivity. It does not seem to pose any challenges for inter-device power control.

Other non-IP Technologies

There are many that can move digital data between devices but that have not gained traction for wide use in A/V systems. It is not likely that these will become widespread — rather new significant technologies are likely to be those new ones that introduce new features or benefits.

Control of A/V devices can be accomplished through buying external hardware that is programmed to manage the functionality and power state of the devices that actually manage the content. These controllers commonly send out commands over infrared (IR) signals, RS-232 ports, 12 V triggers, Ethernet, USB, or radio frequency signals. These technologies are most commonly used on high-end systems that are designed and installed by professional "A/V Integrators," not by ordinary consumers, as they are not in general "plug and play." In addition, they tend to be "brittle" — easily broken — and require a professional any time that a device is added or removed.

Summary

Among all these technologies, HDMI is the most well-developed and the most widely used and so the one of choice to put forward a more comprehensive system.

Digital IP Technologies

The Internet Protocol enables arbitrary communication among any devices on a network — any two can communicate, and use any protocol mutually agreed upon. This is in contrast to the dedicated channels of non-IP technologies which generally provide only peer-to-peer exchange and a single protocol. IP data has additional capabilities such as: passing over many links from source to destination; multi-casting (sent to several devices simultaneously); either the source or destination can be outside the local network; and a given device can have several or many content or other data streams going on simultaneously, with the same or different devices.

In many cases, A/V streams on IP networks also pass over non-IP links at either or both ends. For example, a set-top box may receive an IP video stream and then convert it to HDMI for transport to a TV. Or, a broadcast TV signal may be sent to a "Slingbox," which then sends the stream out on IP to somewhere on the Internet.

Just because a device has an IP connection does not mean that it is used (this is of course true of any technology). Parks Associates recently found[3] that use of IP connections on televisions that have them has "steadily increased, from approximately 40% in 2010 to 56% today". A common usage is streaming services such as Netflix which can be routed through an external IP set-top box, a personal computer, a game console, or a Blu-ray player.

In a non-network context, devices are connected to each other by wires and so there is little question that devices should be communicating with each other for some purpose. However, in a network context, there can be many devices on the network and finding who it is you might want to talk to is a non-trivial issue. This is conventionally divided into two parts: device discovery and service discovery

Device discovery is locating other devices in your vicinity that have one or more common languages that they may speak to each other in. These protocols essentially create the possibility of communication. **Service discovery** is where devices expose to each other the types of functions they can carry out. For example, a printer exposes that it can print. A multi-function device may expose that it can scan, print, fax, email, and provide storage.

In A/V networks, relevant services are providing content (audio and/or video), processing content, storing it, or displaying it. These need to be exposed in protocols.

This section reviews technologies by the layer they cover: first physical layers, then technologies immediately above that, and finally those that are clearly application layer technologies.

Physical Layer Technologies

At the foundation of the Internet is technologies which move IP packets from one device to a second adjacent device or a physical medium. This section covers the physical layer IP technologies used by A/V systems.

Ethernet

Ethernet (IEEE 802.3, ethernetalliance.org) is the foundational wired Internet Protocol technology (ethernetalliance.org). It is inexpensive, flexible, multi-speed, widely available, and reliable. Any building of any complexity will almost certainly include some Ethernet, if not a lot of it. There is even a move to develop a version of Ethernet for use in motor vehicles. With Energy Efficient Ethernet (IEEE 802.3az), Ethernet is particularly effective at matching electricity use to the amount of data actually transmitted (and the CEC had a hand in creating this

[3] Parks Associates, "Smart TVs in a Pay-TV World", July 17, 2012.

technology). An increasing number of A/V devices include an Ethernet interface to enable generic Internet Protocol connectivity among any local IP devices and with the Internet generally.

Wi-Fi

Wi-Fi (IEEE 802.11, wi-fi.org) is the wireless counterpart to Ethernet. Wi-Fi has been highly successful in being deployed and is used in a large number of buildings and individual devices. Early versions of Wi-Fi had maximum data rates that could challenge distribution of A/V signals, particularly when interference or distance reduced transmission below the rated figures. However, in more recent years, Wi-Fi capacity and reliability has significantly increased making it much more suitable for A/V stream distribution.

There is a technology called WiGig, which shifts Wi-Fi technology to frequency bands other than that which traditional Wi-Fi uses. This is now called 802.11ad, and is in development with the standard expected to be finalized at the end of 2012. This enables very high throughput compared to other 802.11 technologies. From the application perspective, WiGig moves IP packets, so this has no fundamental affect on power control. However, there are Protocol Adaptation Layers being defined to tunnel HDMI and Display Port signals efficiently over a WiGig connection, and presumably will enable the same type of peer-to-peer connections that the standard frequency Wi-Fi does (with Wi-Fi Direct, see below). If this layer is fully transparent, then it should have no effect on an HDMI connection, but full transparency is not a certainty.

Some devices (e.g. A/V receivers) use Wi-Fi to enable iOS or Android devices to be remote controls. These are not involved in moving the A/V stream directly, but the receiver should be monitoring the Wi-Fi network while it is asleep to be able to receive commands that should cause it to wake.

MoCA

The Multimedia over Cable Alliance (mocalliance.org) enables transmitting large quantities of IP data over coaxial cable, which is often present in houses to distribute cable TV signals. This repurposes these wires to move digital A/V data. Such coax wiring is usually a tree structure, but MoCA enables ordinary IP network communications among the connected devices over this wiring.

While MoCA was designed with specific application to A/V devices in mind, it is not limited to these devices, and provides generic IP connectivity just like Ethernet and Wi-Fi. That said, the users of MoCA seem to mostly use UPnP for device discovery and management (see below).

Homeplug

Homeplug is a mechanism to put IP packets on the power lines in houses. The HomePlug Alliance has several versions of the standard, including one specifically designed for the high data rates needed by A/V signals. In general, wired communication, including powerline, is more popular in Europe and other places with more masonry walls that are challenging for many wireless technologies.

HDBaseT

HDBaseT (hdbaset.org) is a network technology, derived from Ethernet, that both tunnels HDMI signals over category-5e/6 cabling (commonly used for Ethernet) in either a point-to-point or network configuration (e.g. with a star topology around a switch as a typical Ethernet network is set up). It also enables Ethernet communication along the same path, as well as power delivery similar to (but greater than) what Power-Over-Ethernet provides. Thus, devices with HDBaseT can use either traditional HDMI signaling, generic Internet Protocols, or both.

Summary

The layered design of Internet protocols is intended to isolate complexities and details of particular layers from each other. This means that for the most part, the physical layer that IP protocols travel over do not matter for the application layers above. This is largely true for A/V protocols, and the only issues that do arise, such as quality of service and addressing, are addressed by the technologies in the following section.

Intermediate Layer Technologies

While "pure" physical layer technologies do an enormous amount, over the years, some specific problems or opportunities have arisen that led to technologies which layer closely on top of them. This section describes the three of these most relevant to A/V power control.

Wi-Fi Direct

Traditional Wi-Fi routes all data through a central "access point" (AP) that is the only device with which each end device communicates. Wi-Fi Direct enables one device to communicate directly with a second device within its radio reach. A simple example is to send data from a computer to a printer. There can be an AP in the vicinity or not (and if so, either or both devices can have connections to it) but all communication is directly between the two involved devices. This first requires a process to initiate a session with "Wi-Fi Protected Setup." Devices can also have a group relationship, though one device is the "gate keeper" and determines other devices allowed in the group. In both cases, only one of the devices needs to support Wi-Fi Direct — the others can be an older Wi-Fi device.

Wi-Fi Direct devices can discover each other in two ways. One is if they are both already connected to a single access point, in which case the 802.11z protocol enables them to negotiate switching to a direct connection. The second is a peer-to-peer negotiation in which they use Wi-Fi probes to negotiate establishing a connection. In this case there could be an access point present, or not. This second method is also useful if one or both of the devices do not have authorization to use the access point, which might require a password.

AVB

Audio Video Bridging (AVB) is in the IEEE 802 family of standards that includes Ethernet and Wi-Fi; it is being promoted by the AVnu Alliance (avnu.org). The fundamental issue it addresses is quality-of-service (QOS) — to ensure that an audio/video stream receives the necessary capacity on network links and a minimum of delay, so that the stream which transits a local area network is complete and timely. It works for both Ethernet and Wi-Fi, which are also in the 802 family. It is designed to have the sophistication and quality needed for professional audio installations (e.g. theaters), but is simple enough to be used in products designed for ordinary home use. AVB should be transparent to upper layers and so should not affect power control.

IEEE 1905.1

This standard, on "Convergent Digital Home Network for Heterogeneous Technologies", enables easier integration of different physical layers to form a network that is more homogeneous from the perspective of each device. This makes the higher layer networking simpler and more uniform, removing problems that might otherwise occur for users. The standard claims "energy management" as a benefit saying that it can "optimiz[e] network power usage across different technologies". It does clearly expose the power state of individual network interfaces of devices, and enables setting them. It is unclear if it exposes the power state of the entire device.

Summary

These intermediate layers simply provide a more seamless and efficient way to access the capabilities of the underlying physical layers, but don't change what is ultimately done. The only apparent impact for A/V control is to ensure that devices can be asleep, and/or already connected to a wireless access point, and respond to appropriate Wi-Fi Direct queries.

Service Discovery Protocols

As noted above, service discovery is how devices know that they have some purpose for a communication to occur, namely, that one device can provide a particular service to another. Two such open protocols are commonly used with A/V devices. Some devices implement both of these[4].

mDNS (Bonjour)

Apple created the Bonjour technology so that its own products could interoperate on a network, but ultimately donated it to the public domain in the form of mDNS (multicast Domain Name System). For a local network, it enables both device discovery and service discovery. Core Apple technologies such as iTunes use it to move A/V content from one device to another, whether the underlying hardware is made by Apple or not. For example, a Windows PC with iTunes on it has mDNS installed.

A key then is to ensure that devices stay on the network for mDNS purposes while asleep. Apple and other companies have grappled with this general issue for several years already, and the CEC previously contributed to one, the "network proxying" technology (Ecma, 2010). The Ecma standard (Ecma-393) covers mDNS.

UPnP

UPnP (upnp.org) is a widely-referenced mechanism for service discovery and device control, including for A/V devices, particularly by the Digital Living Network Alliance (dlna.org). It is very complicated, which brings with it disadvantages for implementation.

A NVidia document on Miracast (NVidia, 2012) states that "…DLNA [is] plagued with interoperability issues that have limited [its] traction" (DLNA is based critically on UPnP). UPnP is clearly problematic, but also has a large base of installed products and so must be part of the solution. UPnP has dealt with the issue of products being in low power modes several times to date, considering proxy approaches.

The RVU protocol (rvualliance.org) is layered on top of UPnP to provide greater interoperability and a consistent user interface.

Remote Display Protocols

An increasingly common situation people find themselves in is to have media (usually video but also sometimes only audio) on one device and want to display it on a second local device over a wireless connection. Several protocols have arisen to address this. Note that while we think of visual entities for the word "display", in principle a "display" can be solely audio — a speaker is a display.

Airplay

Airplay is an Apple technology for distributing audio/video content over an IP network. An example application is an A/V receiver that can use an Airplay device (computer, tablet, or music player) as a source of A/V content, to then relay that to a connected TV (likely over HDMI).

---

[4] An example is the Russound DMS-3.1 Digital Media Streamer.

Airplay is built on several open standards, but integrated in a way that has not been standardized except through Apple. The first step is for devices to discover each other on the network, and determine what services they have and may want to utilize; the mDNS protocol is used for this (see above). Different protocols separately manage audio and video streams (Airplay was originally an audio-only service with a different name). Control of the audio stream is done with the Digital Audio Control Protocol (for play, pause, stop, etc.). Audio is transmitted using the Real Time Streaming Protocol (RTSP).

Video is controlled with a series of HTTP requests, and video is understood to be delivery of an item of known and finite length. The specific commands sent over HTTP are Airplay-specific. A separate mechanism is used for "screen mirroring", e.g. to duplicate the content of one display on a second display over the network; mirroring has no specific length or end.

Key for this project is for a device to maintain its presence in the mDNS system even while asleep. This author owns an Apple TV device which maintains its presence on the network as an Airplay device while asleep, while consuming only 1.8 W.

WiDi

WiDi is an adaptation of the Wi-Fi standard created by Intel (Intel, 2012). Some new TVs and computer displays and some new source devices (computers, a game console, and an Internet set-top box) already have WiDi built in. However, since the Wi-Fi Alliance has created Miracast to accomplish essentially the same functionality, it is unclear if WiDi has any significant future.

Miracast

Miracast is the name for a Wi-Fi technology endorsed by the Wi-Fi Alliance to transmit video data from one device to another. It is based on Wi-Fi Direct, which enables peer-to-peer Wi-Fi connectivity, in contrast to the usual use of Wi-Fi in which all traffic flows only to and from a central "access point." As part of the Wi-Fi Alliance, Miracast is characterized as being open in contrast to WiDi and Airplay [NVIDIA], controlled by Intel and Apple respectively, and to other systems that are entirely proprietary.

In Miracast, devices first connect via Wi-Fi Direct, and then use a Miracast-specific protocol to accomplish service discovery. Wi-Fi Alliance standards are almost always confidential to the organization so that the details of this mechanism are not known.

Other Technologies

Some other technologies are in the A/V space but not specifically a communications technology.

The UltraViolet technology maintains libraries of digital assets (uvvu.com), and does not directly interact with end-use devices. Rather, a provider of content, such as NetFlix, uses UltraViolet to identify content owned by individuals.

The Smart TV Alliance (smarttv-alliance.org) recently formed and aims to create a platform environment for TV "apps" that are independent of more proprietary systems such as iOS or Android. They will use HTML5 as the core technology. At this time, only two TV manufacturers are on the board of directors so whether it will take off remains to be seen.

A/V Streams

From the beginning of audio/video systems[5], media content was strictly a 1:1 relationship, with a single source and a single sink of data involved. The first change to this was the introduction of an A/V Receiver, which sends sound to a speaker system (rather than use internal TV speakers), and enabled higher quality sound, multiple speakers, etc. The source devices are typically all connected to the A/V Receiver, which in turn then has a single connection to the television. However, even with this complication, there was still a single stream of content and all the video was displayed by a single television.

For several decades, some products have included a "Picture-in-picture" capability in which the video from one device is inserted in a small box in the video of a the main display source. This enables monitoring a sports event while the TV news is on, for example. However, in recent years, the number of types of sources and applications for TVs has continually increased. For example, some have cameras and microphones and have the Skype application loaded for videoconferencing. Many TVs can stream content from the Internet. We have protocols (described above) to mirror the display of a computer onto a TV, and should expect the reverse to be enabled as well. Security camera systems commonly combine the output of a few or many cameras onto a single display. Video streams can be split, to send the same content to multiple displays in the same or different rooms, or (as with the Slingbox product) arbitrarily across the Internet. In addition, audio doesn't have to follow the same path as the video. Usually it originates in the same place, though can be displayed through a separate communications path from the video. Some video doesn't have audio, and some content is audio-only.

Streams are becoming more complex, with branches on either end, multiple sources, and/or multiple sinks. Even simple radio is becoming more complicated, with increasing use of "Internet Radio" rather than the traditional broadcast sources. As capabilities grow, human users as well as devices need to be able to conceptualize and communicate the nature of A/V streams. Protocols will need to be adapted and expanded to cover the new capabilities. For example, the DisplayPort technology has the concept of a "concentrator", with multiple inputs and one output, and a "replicator", with one input but multiple outputs.

A video stream may get altered on its path without losing its fundamental character. For example, the addition of a small picture-in-picture (PIP) to a movie is best thought of as the additional stream ending in the device which does the PIP, with the movie stream possibly continuing on past the PIP device. Other examples of stream alteration are the addition of captions, or changing the resolution or aspect ratio of the image. On the other hand, when 16 security camera images are combined, they are best thought of as all ending in the combining device with a new stream emerging.

Streams are becoming more dynamic. Display devices (and even intermediate devices) may be added to a stream while it is active, or a device may be dropped from a stream that was essential to it originally. For example, a TV may be playing a movie, then mirrored to a second TV in a different room, then the first TV switched to a different source. Depending on the context, the user might want the second TV to also switch, or alternately, have it keep displaying the movie. The same could occur with audio. Network connections make this all easier, with their arbitrary connectedness. Peer-to-peer A/V links are more limited in the scenarios they can implement.

It is not apparent that there is any other example of a technology which requires such sophisticated coordination of multiple devices in a generic user context. Even in web browsing,

[5] A standalone device does not implement an A/V stream but it is not a "system".

from the user perspective they have just a 1:1 relationship between a URL and a web window (or tab within window) — there is no further device complexity.

Regardless of the stream topology, many streams use different physical layers on different links, including mixtures of IP and non-IP technologies. This necessarily means that some devices implement both. Also, some technologies enable transmission over both, and IP can supplement or replace non-IP communication.

Control Paradigms

The power control topic covers two parallel mechanisms — functional state and power state. These are linked in that a device must be powered up to perform an active function, and can be powered down if there is no active function to carry out. Devices and protocols variously include some of each of these, with HDMI providing the best example. HDMI includes commands to change the power state of a device — up or down — as well as commands to change the functional state of a stream, specifically, what devices are selected to be part of a stream. Powering up a source device will typically cause it to start sending data, or, selecting a device will cause it to power up. Similarly, powering down a device will shorten or terminate a stream, and terminating a stream gives an indication that a device should power down (possibly after a modest delay time). There is also the case where a communications link is unintentionally lost, which might be temporary, or indefinite. Devices need to accommodate both situations, reconnecting the link when possible, or after a suitable time, determining that it is more permanent, and going to sleep. Pausing a stream is a similar situation.

All power control was initially manual, physically located on each device, and later, with remote control commands. Remote controls first could power on the parent device only, then some remotes could also send power commands to other devices. Finally, "one touch play" and "all off" capabilities were added to remotes that accomplished several actions at the same time, both functional and power control. A "one touch play" command tied to a DVD, for example, would ensure that both the TV and DVD player were powered up, power up any intermediate device (like an A/V receiver), and select the correct input on the TV and any intermediate devices. This is essentially creating a stream.

Thus, we are in the midst of moving from a solely power control world to a mostly stream-oriented world. As devices increasingly can have multiple functionalities (possibly unrelated), as PCs do, the distinction between power state and functional state becomes ever greater. Network connectivity enables arbitrary connections among devices, unlike the peer-to-peer connectivity of traditional A/V communications. While stream orientation will likely dominate, there will remain a role for more manual power control for the foreseeable future.

Another issue which will be an increasing feature of A/V systems is streams which pass across multiple A/V technologies on their way from source to sink. Thus, knowledge of the entirety of a stream needs to be greater than a single technology, or at least bridged at a device which is the transition point. In either case, the semantics of stream function and control then need to be similar across technologies to enable smooth performance and functionality.

Conclusions and Next Steps

Communications technologies fall into four basic categories for our purposes: One-way (mostly analog), HDMI, Other Non-IP digital, and Network (IP). Each technology has inherent characteristics which govern how functional and power states are managed between devices. HDMI and IP networking likely hold the key to successfully evolving the whole range of standards. Even when the other technologies are used, it is likely that these two will be core parts of the audio/video system.

For one-way communication, sending devices can have time-outs to stop sending data after a period of time of no indication that the stream is being consumed. Receiving devices can monitor incoming streams and power down when none is present, but in many cases, they should continue to monitor one or more inputs in case data reappears (subwoofers are a basic example of this behavior). One final mechanism available for certain interfaces is detection of connected cables. Some devices do this, and take action such as turning off internal speakers when a jack is inserted into the headphone output. There may be some cost to this hardware capability, and some analog interfaces do not lend themselves to this capability. For this project, identifying suitable standard default time-outs is the only apparent outcome to work toward.

In November, 2011, the HDMI standard recently transitioned from being open to being confidential to the members of the new HDMI Forum. We are in the process of trying to arrange a liaison agreement to be able exchange confidential information but that is still pending. Having such an important standard not be in the public domain raises questions for how to enable interoperability and gain energy savings, but it may be that enough key aspects of changes to the standard could be made public to remove this as an issue.

As the most widely used digital A/V technology today, and the most developed with respect to power control, it makes sense to build as much as possible on the principles embodied in HDMI/CEC. However, there needs to be extensions to what CEC does today. The HDMI specification already includes elements — for example, it is noted that a "system standby" command for all devices to power down (most likely emanating from the TV) should be ignored by a connected device that is in the process of recording something (whether that content was being displayed on the TV, or not). Also, an HDMI signal might usefully pass through a TV and on to other devices even if the TV itself is not displaying audio or video.

After the approach to HDMI and general IP technologies becomes clear, those conclusions can be applied to other non-IP technologies.

Many IP technologies are at a sufficiently early stage of development that the types of usage scenarios of concern have yet to be sufficiently explored and demonstrated. What is clear for IP is that service discovery needs to work for devices that are asleep. This remains an outstanding issue for UPnP. Airplay and UPnP seem to be the best places to engage A/V power control in the IP context.

What is needed next is a larger example set of use cases for power control and stream management, and a system for describing this and conceptually managing it. This could include the concept of a "sleeping stream" proposed in the ACEEE paper prepared as part of this project. That is the subject of the next deliverable for this project.

Acknowledgement

Many thanks to Gari Kloss for her invaluable contributions.

References

Ecma International, *ECMA-393, proxZzzy for sleeping hosts, Edition 1*, Accessed September 3, 2012, http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-393.pdf

HDMI, *High-Definition Multimedia Interface Specification*, Version 1.1, May 20, 2004.

HDMI, *CEC Implementation Guideline*, Version 0.9, January 2009.

Intel, *How to Connect Laptop to TV with Intel® Wireless Display (WiDi)*, accessed Sept 3, 2012, http://www.intel.com/content/www/us/en/architecture-and-technology/intel-wireless-display.html

Intel, *Thunderbolt Technology: The Fastest Connection to your PC*, 2012, Accessed September 3, 2012, http://intel.com/content/dam/doc/technology-brief/thunderbolt-technology-brief.pdf

NVIDIA, *NVIDIA® Miracast Wireless Display Architecture*, Version 1.0, July 24, 2012.

VESA, Video Electronics Standards Association, *Display Power Management Signaling - DPMS 1.0*, 1993.

VESA, *VESA DisplayPort Standard*, Version 1, Revision 1a January 11, 2008.

Vasseur, Clément, *Unofficial AirPlay Protocol Specification*, accessed September 5, 2012, http://nto.github.com/AirPlay.html, March 2012.

Wi-Fi Alliance, *Wi-Fi Certified Wi-Fi Direct*, October 2010.

WirelessHD, *WirelessHD Specification Overview*, Version 1.1, May 2010.